

LA-UR-19-25653

Approved for public release; distribution is unlimited.

Title: Data Analytics for Nonproliferation Applications

Author(s): Ruch, Marc Lavi

Intended for: IAEA Visit to LANL on 6/19/2019

Issued: 2019-06-18

Disclaimer:

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by Triad National Security, LLC for the National Nuclear Security Administration of U.S. Department of Energy under contract 89233218CNA000001. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Data Analytics for Nonproliferation Applications

IAEA Department of Safeguards Visit to LANL
June 19, 2019

Marc L. Ruch

Safeguards Science & Technology (NEN-1)
Nuclear Engineering & Nonproliferation Division



Managed by Triad National Security, LLC for the U.S. Department of Energy's NNSA

Data Analytics: What are we even talking about?

As a phrase, not in the dictionary (yet things like “machine learning”, “neural network”, and “data mining” are...)

Analytics

- Oxford: “*The systematic computational analysis of data or statistics.*”
- Merriam-Webster: “*the method of logical analysis*”

Exceedingly broad definitions: “every sane thought is analytics”

So, I took a survey of the LANL Safeguards Science & Technology group:

“Approaches to analyzing large sets of data with specific computer-aided tools to extract meaningful information that would be too time-consuming or complex for human analysis.”

“Analysis of a big data set to get some information about the problem that one would not be able to get from a simpler data set or a single data set.”

Positioning for the Future

Traditional capabilities in instrumentation, modeling and simulation, and fuel cycle expertise are enduring needs, but where else could we be focusing R&D efforts to prepare for the future of safeguards?

Machine Learning & Automation

Artificial Intelligence
Data Fusion
Anomaly Detection
Predictive Modeling
Data Visualization

Distributed & Persistent Sensing

Remote Monitoring
IoT Sensors
Wireless Technology
Authentication
Data Security

Non-Traditional Data Streams

Geospatial Data
Satellite Imagery
Publications
Open Source
Multimedia

Data analytics will play a central role in each of those areas!

Advanced Sensing

Wireless Sensor Networks

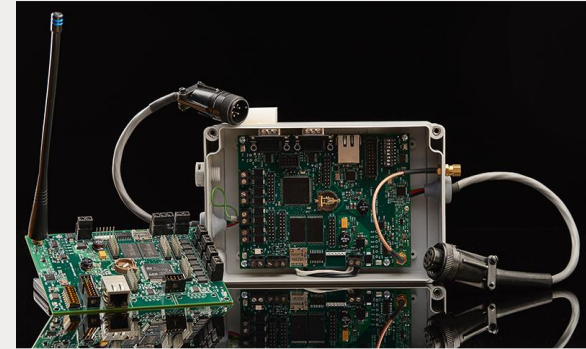
POC: Janette Frigo

Long-range wireless sensor network for monitoring stormwater runoff in remote areas

- Self-forming mesh network scales to ~200 nodes
- Sensors can operate continuously for 1-5 years
- Capable of onboard data processing prior to transmission



The research team configures a stormwater runoff sensor node in the field

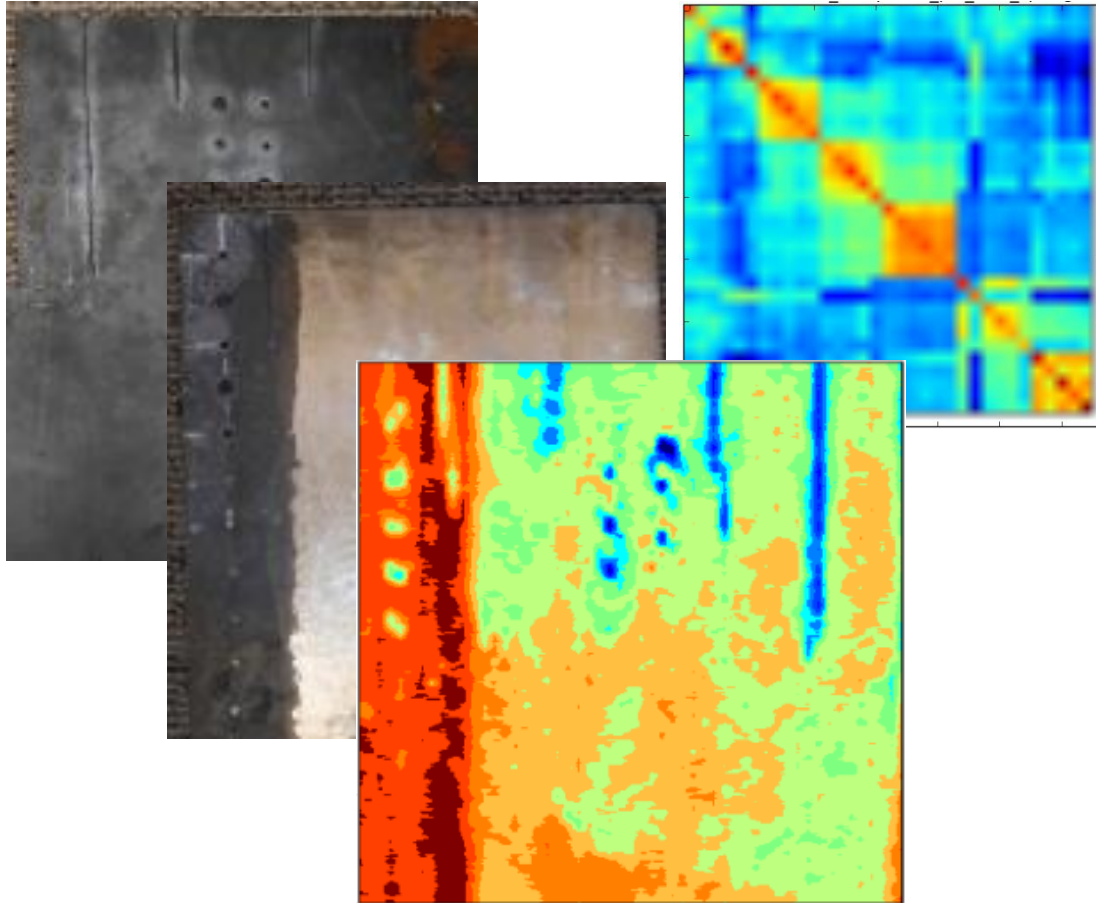


We are leveraging the sensing capabilities developed in the stormwater project and applying them to problems in nuclear facility monitoring for safeguards & security



Barkhausen Noise: Intrinsic Fingerprint for Ferromagnetic Items

POC: David Mascareñas



Barkhausen Noise provides information concerning a ferrous material down to the microstructural level

A unique fingerprint of ferromagnetic materials

Provides evidence of tampering and clandestine repairs

Measurable using commercially available equipment

Facility Control Shield

POC: Vlad Henzl

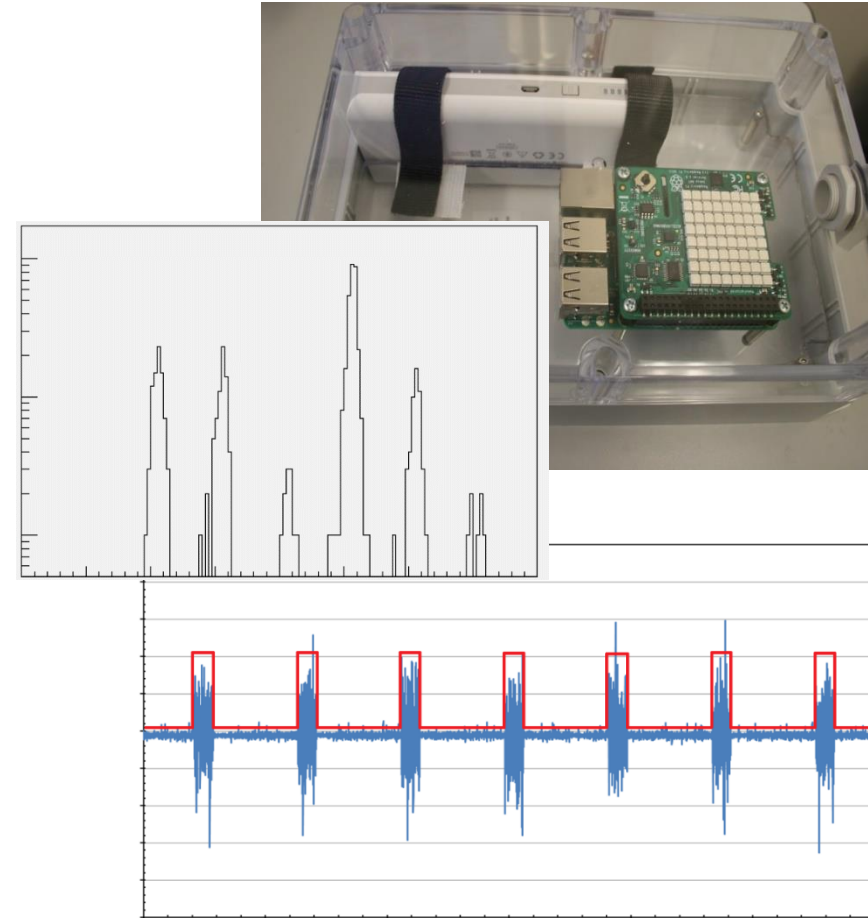
Uses small distributed sensors to verify a supervisory control and data acquisition to prevent of cyber hijacking of facility

Sensor node cost: \$74

11 sensors installed in a facility in 1 hour

Vibrational and temperature data used to establish “cadence of facility,” when unit ops in use

Plans for analysis in *unsupervised mode*



Disparate Data Integration for Improved Safeguards Verification

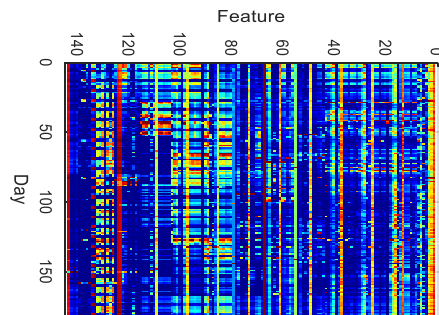
POC: Vlad Henzl

Problem: Vast amounts of safeguards data nearly impossible to analyze only by human experts, especially if many data streams are incompatible with one another

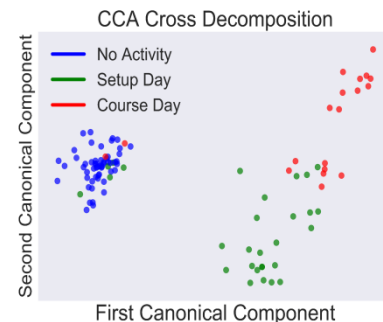
Solution: Develop automated methods and approaches to integrate large, heterogeneous data sets for *patterns-of-life analysis*



Disparate data collection
(nuclear facility testbed)



Data feature extraction
(1 data stream → multiple
data features)

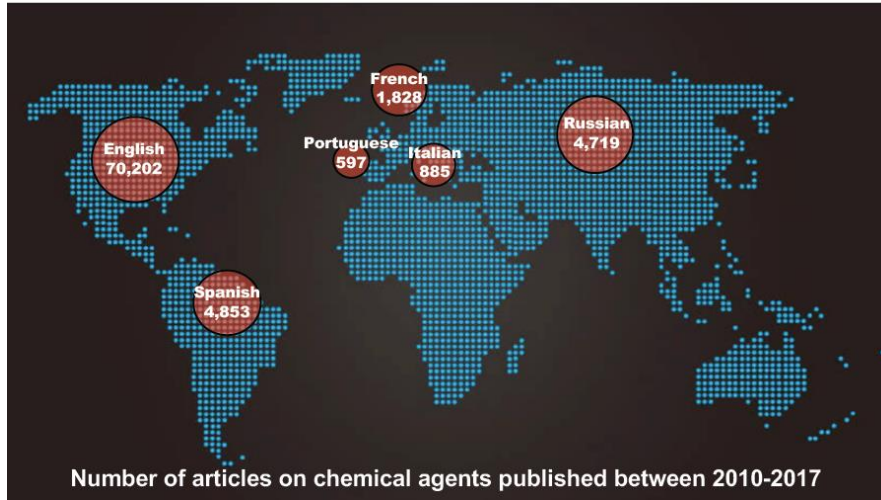


Validated Models
(for decision support)

Open Source & Geospatial Data

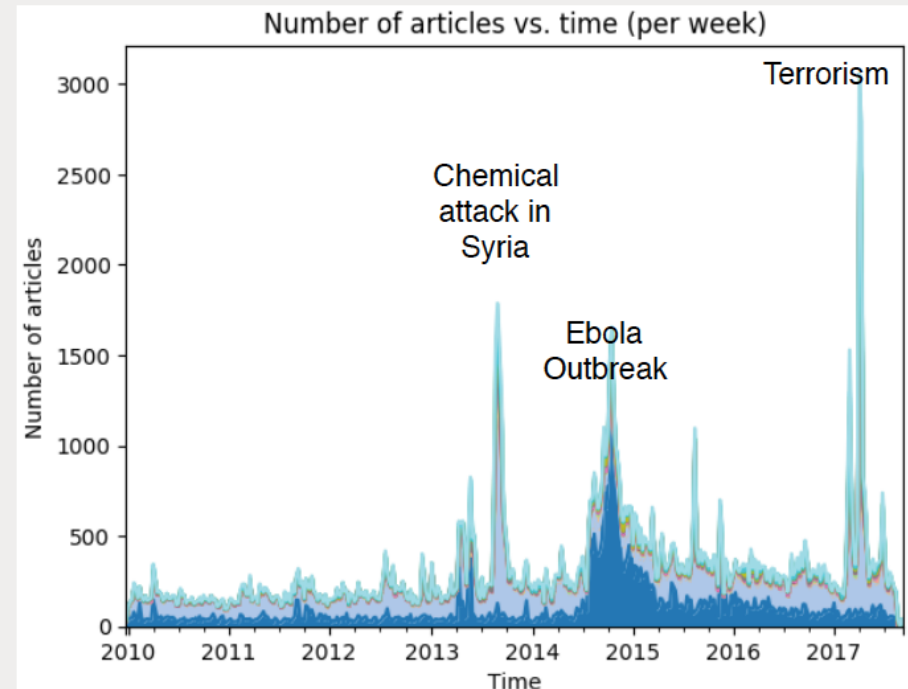
Trend Analysis using Dow Jones Data News Analytics

POCs: Sara Del Valle & Geoffrey Fairchild



Los Alamos partnered with Dow Jones to pilot their Data, News, and Analytics (D/DNA) platform to determine its potential to inform national security programs

Time series analysis and topic modeling applied to chemical and biological mentions in news articles



Processing Satellite Imagery at Scale

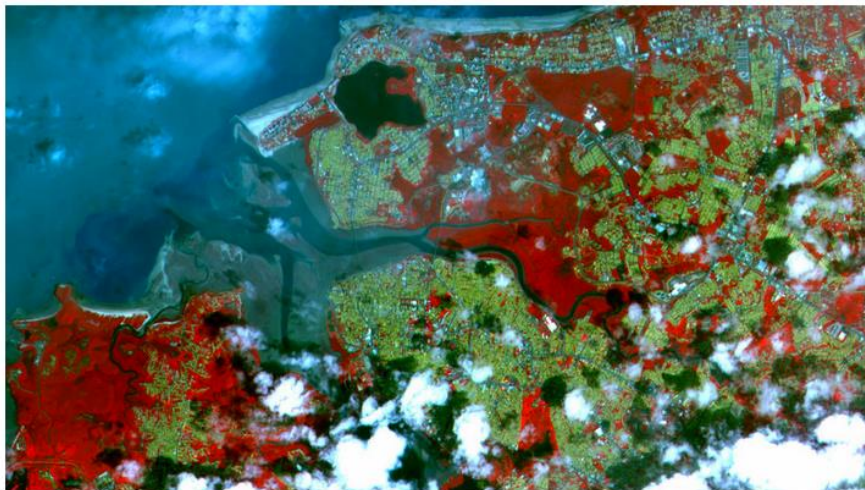
POCs: Nick Generous & Amanda Ziemann

FAST COMPANY

06.13.18 | ROBOT REVOLUTION

How satellite imagery could combat infectious diseases around the world

The Los Alamos National Lab has worked with Descartes Labs to come up with systems for analyzing on-the-ground conditions in Brazil in order to forecast dengue. The hope is to expand the work around the world.



The availability of open source satellite imagery is growing at an unprecedented rate

Los Alamos is using commercial resources to build a disease forecasting tool that combines heterogeneous data streams and expert-driven models

Data processing must be automated to make it scalable to a global level

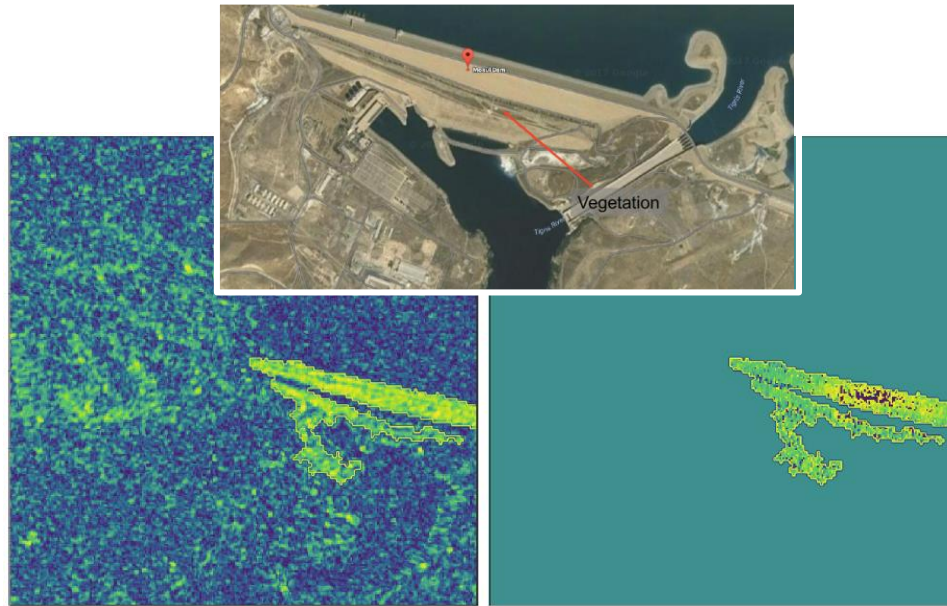
This is truly a big data problem

- Available: petabytes of data
- Pulling: 13 terabytes of data
- Processing on the cloud: 3 virtual machines

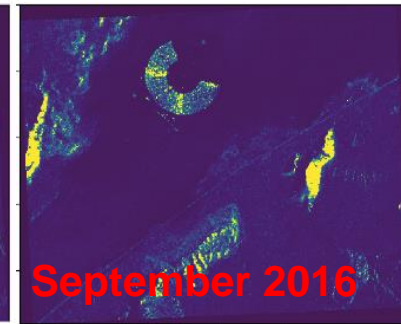
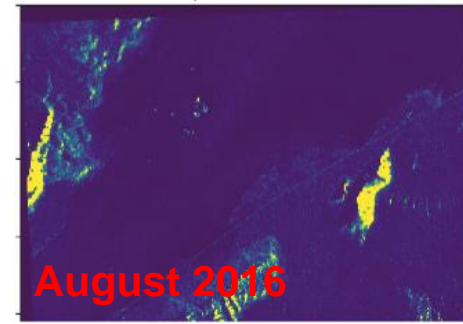
Automated Analysis of Sentinel-1 SAR Data

POC: Amanda Ziemann

Closing the gap between the amount of data we collect and the amount of data we use in time to impact decisions



Subsidence detection applied to the failing Mosul Dam using a machine learning approach



Anomalous change detection applied to SAR images before and during the Burning Man

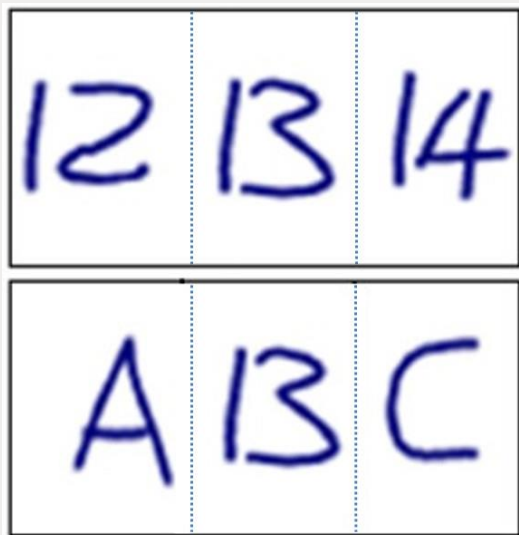
Festival

Sense-making

Context-Aware Computing

POC: Alexei Skurikhin

Context disambiguates information derived from data collections



Recognition of the central object as “B” or “13” is **primed by context** (i.e., how the adjacent objects are recognized)

Machine learning based on probabilistic graphical modeling to capture and exploit contextual dependencies in multi-modal data for recognition of complex object or events

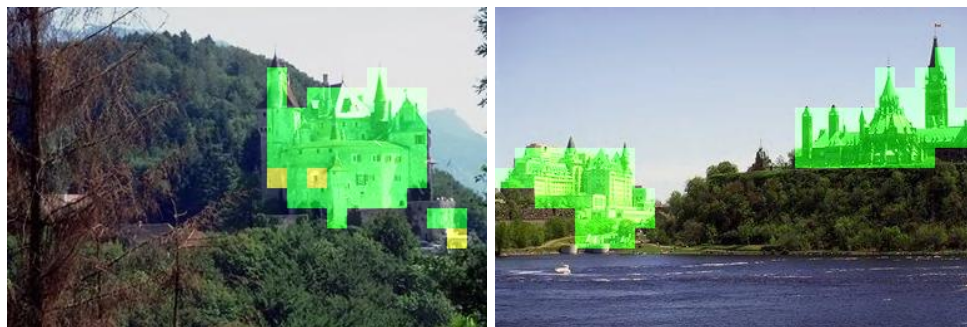


Illustration of the detection of image regions with man-made structures

Green = true positives

Yellow = false negatives

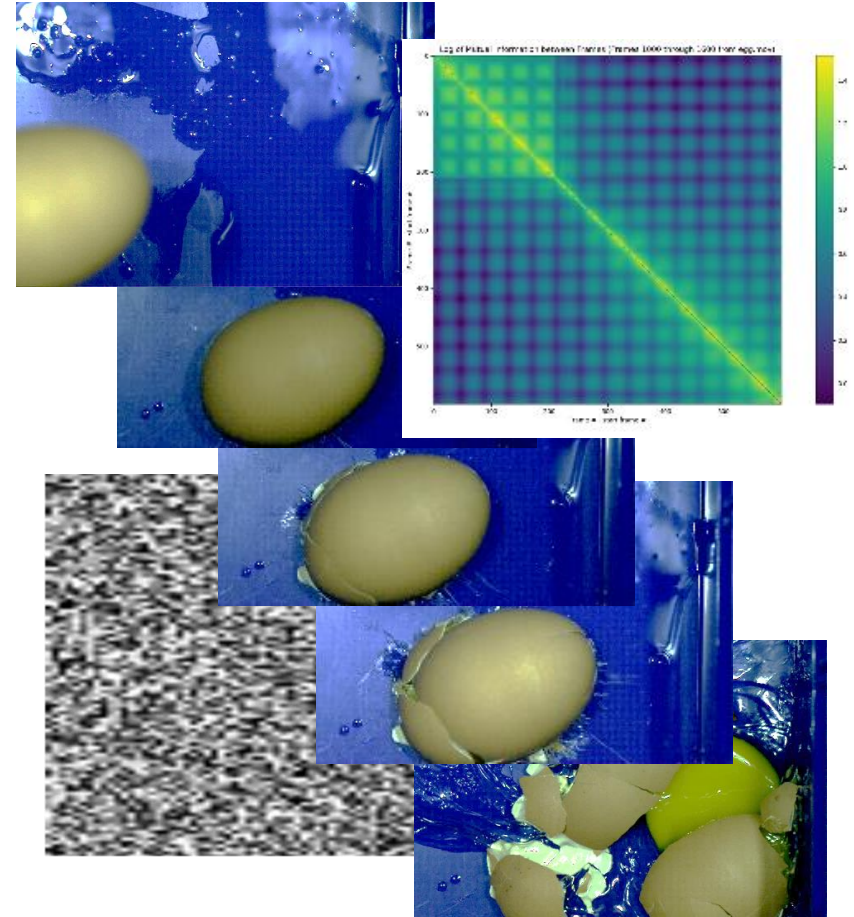
Privacy-Preserving Disassembly Verification

POC: David Mascareñas

A concept to preserve privacy while detecting the occurrence of disassembly in an image feed

Mutual Information between frames shows a distinct change at the time of disassembly

Privacy can be further enhanced by sampling a subset of randomly shuffled pixels for analysis → same distinct features in Mutual Information seen at time of disassembly



Omniscient

POC: Marc Ruch

Safeguards data review tool

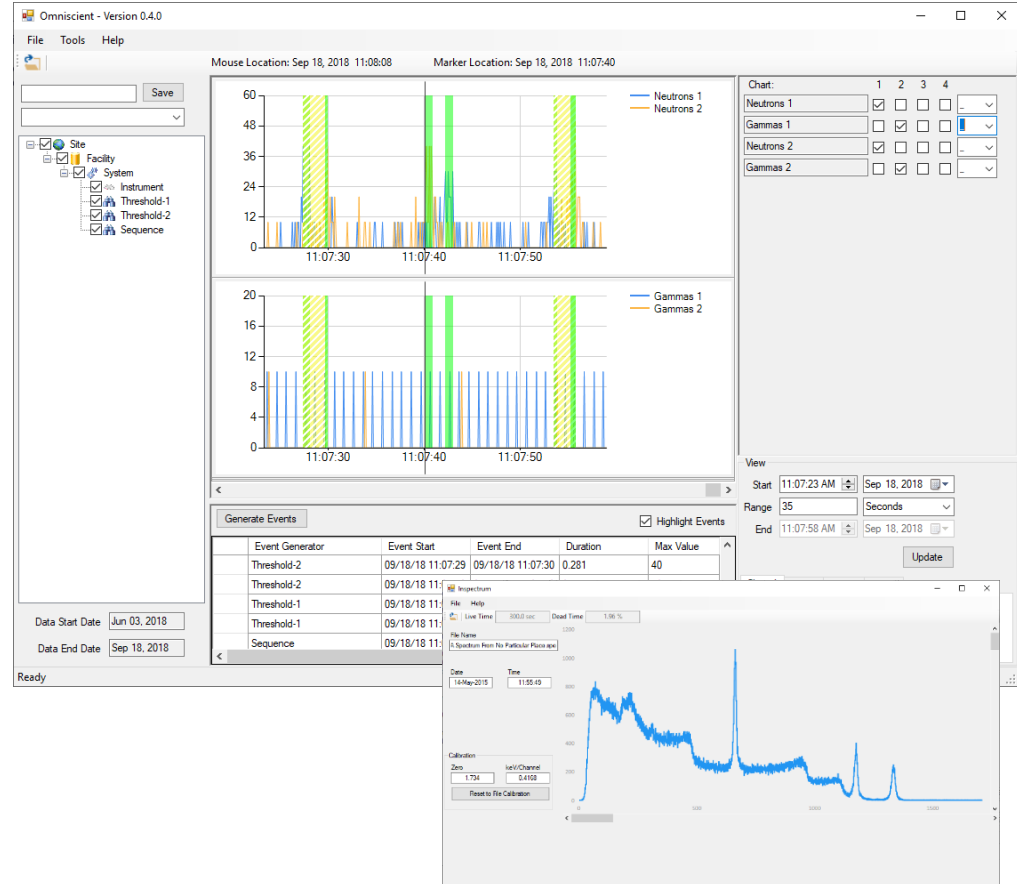
Ingests a variety of data

- Instrument data: gamma spectra, shift register data, images...
- Other data: state of health, declarations...

Displays data at different levels of analysis in a common interface

Automatically highlights and lists events of interest

Both a traditional review tool and a safeguards “data analytics playground”



Novel Technologies

Pattern-Recognition for Automated Co-adding Microcalorimeter Pixels

POC: Mark Croce

Microcalorimetry generates high-fidelity non-destructive safeguards measurements

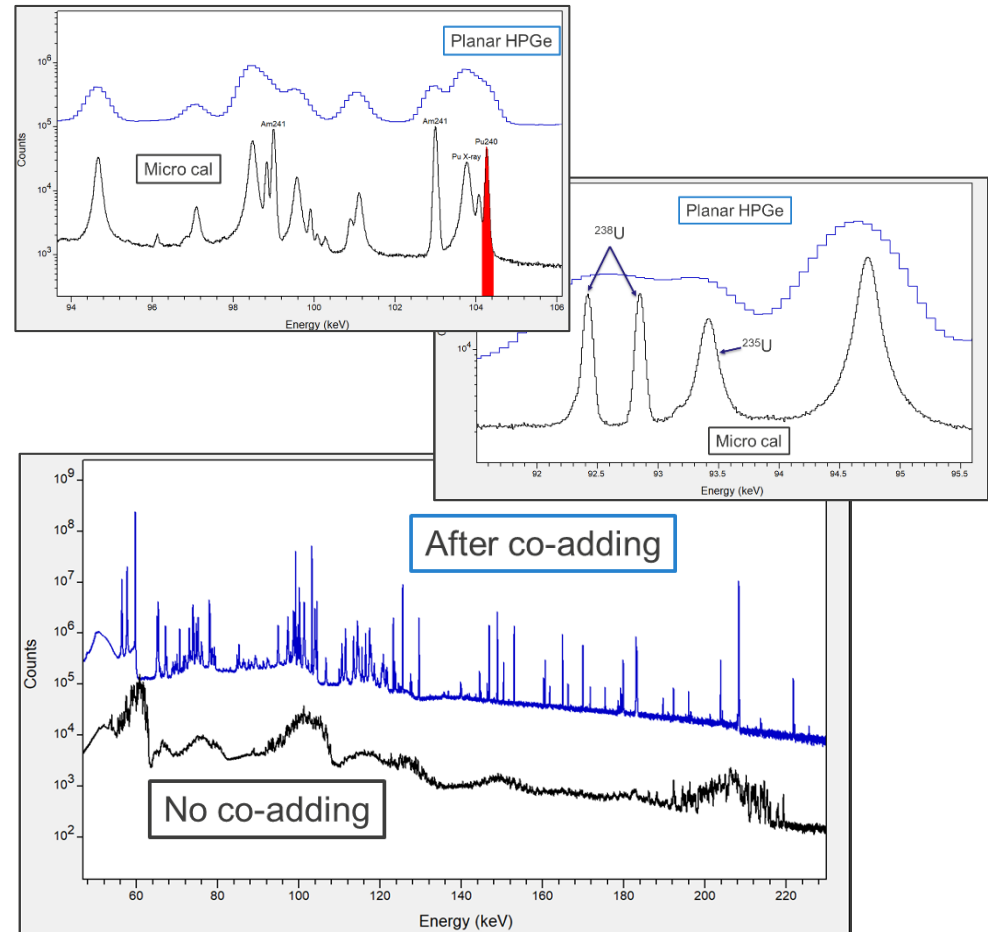
Pixel response highly non-linear

Automated pattern recognition co-adds the 256+ pixel spectra together

Achieves < 10 eV resolution

NDA mapping of uranium *chemical* forms (hyperspectral x-ray imaging)

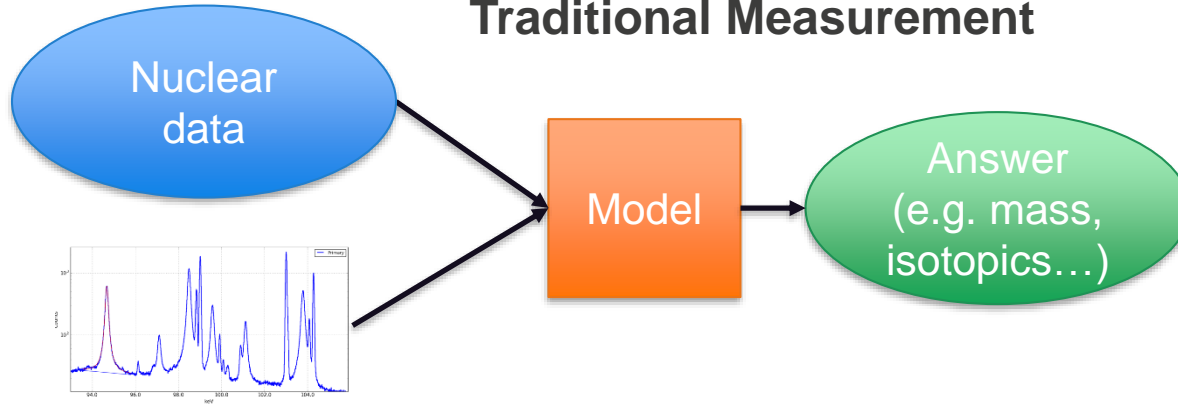
Traditional U/Pu/Mox isotopics



Utilizing Analytics to Improve Safeguards Nuclear Data

POC: Duc Vo

Traditional Measurement

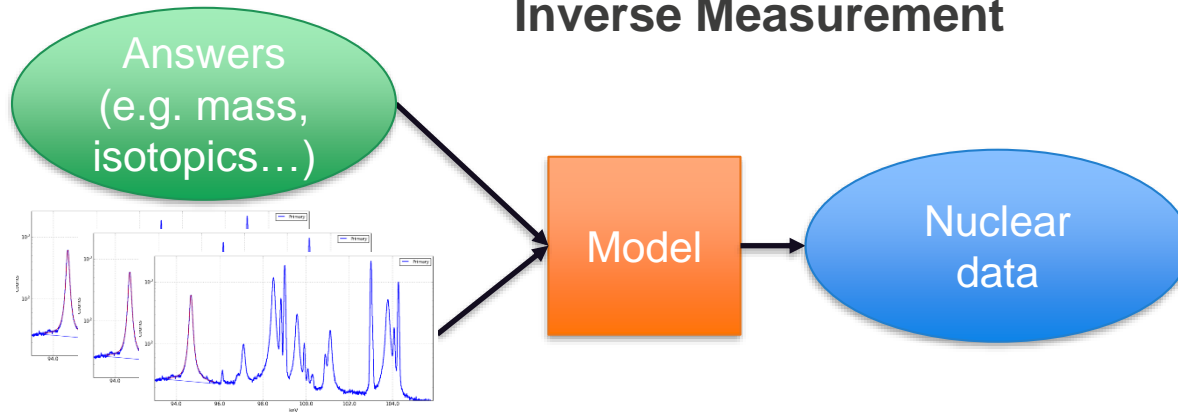


Uses thousands of LANL measurements as training data

Results incorporated into **FRAM 6.1**

- To be licensed to commercial vendors and IAEA in FY 2020

Inverse Measurement



More precise data → **< 0.5% TMU for NDA measurements using microcalorimetry**

Beyond Moore's Law

POC: Scott Pakin

WIRED

KLINT FINLEY BUSINESS 01.11.17 1:00 PM

QUANTUM COMPUTING IS REAL, AND D-WAVE JUST OPEN-SOURCED IT



qubits for All

Qbsolv joins a small but growing pool of tools for would-be quantum computer programmers. Last year Scott Pakin of Los Alamos National Laboratory—and one of Qbsolv's first users—released another free tool called [Qasm](#), which also eases the burden of writing code for D-Wave machines by freeing developers from having to worry about addressing the underlying hardware. The goal, Ewald says, is to kickstart a quantum computing software tools ecosystem and foster a community of developers working on quantum computing problems. In recent years, open



Conclusions

The world is at an inflection point where our ability to collect data now far outpaces our ability to make use of it

Data analytics is a force multiplier for complex problems

The safeguards group at Los Alamos has a focused effort to engage with data analytics experts across disciplines to identify technologies that may provide transformational capabilities for the safeguards community